The Astrophysical Journal, 895:116 (12pp), 2020 June 1 https://doi.org/10.3847/1538-4357/ab8f97
© 2020. The American Astronomical Society. All rights reserved.

# The Ionizing Photon Production Efficiency ($\xi_{\rm ion}$) of Lensed Dwarf Galaxies at $z \sim 2$*

Najmeh Emami[1], Brian Siana[1], Anahita Alavi[2], Timothy Gburek[1], William R. Freeman[1], Johan Richard[3], Daniel R. Weisz[4], and Daniel P. Stark[5]
[1] Department of Physics and Astronomy, University of California Riverside, Riverside, CA 92521, USA; najmeh.emami@email.ucr.edu
[2] Infrared Processing and Analysis Center, Caltech, Pasadena, CA 91125, USA
[3] Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France
[4] Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA
[5] Steward Observatory, University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721 USA
*Received 2019 December 11; revised 2020 April 29; accepted 2020 May 1; published 2020 June 3*

## Abstract

We measure the ionizing photon production efficiency ($\xi_{\rm ion}$) of low-mass galaxies ($10^{7.8}$–$10^{9.8}\,M_{\odot}$) at $1.4 < z < 2.7$ to better understand the contribution of dwarf galaxies to the ionizing background and reionization. We target galaxies that are magnified by strong-lensing galaxy clusters and use Keck/MOSFIRE to measure nebular emission-line fluxes and Hubble Space Telescope to measure the rest-UV and rest-optical photometry. We present two methods of stacking. First, we take the average of the log of H$\alpha$-to-UV luminosity ratios ($L_{\rm H\alpha}/L_{\rm UV}$) of galaxies to determine the standard $\log(\xi_{\rm ion})$. Second, we take the logarithm of the total $L_{\rm H\alpha}$ over the total $L_{\rm UV}$. We prefer the latter, as it provides the total ionizing UV luminosity density of galaxies when multiplied by the nonionizing UV luminosity density. $\log(\xi_{\rm ion})$ calculated from the second method is ∼0.2 dex higher than the first method. We do not find any strong dependence between $\log(\xi_{\rm ion})$ and stellar mass, far-UV magnitude ($M_{\rm UV}$), or UV spectral slope ($\beta$). We report a value of $\log(\xi_{\rm ion}) \sim 25.47 \pm 0.09$ for our UV-complete sample ($-22 < M_{\rm UV} < -17.3$) and ∼25.37 ± 0.11 for our mass-complete sample ($7.8 < \log(M_*) < 9.8$). These values are consistent with measurements of more massive, more luminous galaxies in other high-redshift studies that use the same stacking technique. Our $\log(\xi_{\rm ion})$ is 0.2–0.3 dex higher than low-redshift galaxies of similar mass, indicating an evolution in the stellar properties, possibly due to metallicity or age. We also find a correlation between $\log(\xi_{\rm ion})$ and the equivalent widths of H$\alpha$ and [O III] $\lambda$5007 fluxes, confirming that these equivalent widths can be used to estimate $\xi_{\rm ion}$.

*Unified Astronomy Thesaurus concepts:* Dwarf galaxies (416); Galaxy evolution (594); High-redshift galaxies (734)

## 1. Introduction

Many studies have demonstrated that by $z \sim 6$ the neutral hydrogen in the intergalactic medium (IGM) was mostly ionized (Becker et al. 2001; Fan et al. 2006; McGreer et al. 2015). What is not well understood is what are the sources that ionized the universe and provided the IGM thereafter (Fan et al. 2001; Somerville et al. 2003; Madau et al. 2004; Bouwens et al. 2015b). In fact, it is not clear whether the galaxies that we have detected at high redshift are capable of ionizing the IGM (Robertson et al. 2015; Finkelstein et al. 2019; Naidu et al. 2020; Mason et al. 2019). In order to determine this, we need to know the rate of ionizing photons emitted into the IGM as a function of redshift (often referred to as $\Gamma(z)$). In order to calculate $\Gamma(z)$, three quantities must be known:

$$\Gamma = \int L\Phi(L)\xi_{\rm ion}(L)f_{\rm esc}(L)dL. \quad (1)$$

The first quantity is the luminosity function of galaxies, $\Phi(L)$, which is typically measured in the nonionizing ultraviolet (UV), as it is relatively easy to detect galaxies at those wavelengths at high redshift. If the UV luminosity function is integrated, it gives the total nonionizing UV luminosity density at a given redshift. The second quantity that is needed is a conversion from the nonionizing UV luminosity density to ionizing UV luminosity density. This conversion is often referred to as $\xi_{\rm ion}$ and is defined as the rate of ionizing photon production normalized by the nonionizing UV luminosity density (in $f_\nu$). The third necessary quantity is the fraction of ionizing photons that escape into the IGM, referred to as the escape fraction, $f_{\rm esc}$. Of course, all of these quantities can vary with luminosity, stellar mass, age, and metallicity.

Many studies have constrained the luminosity functions (Bouwens et al. 2006, 2007; Reddy & Steidel 2009; Oesch et al. 2013; Alavi et al. 2014; Bouwens et al. 2015a; Mehta et al. 2017) and escape fractions (Inoue et al. 2006; Siana et al. 2007; Wise & Cen 2009; Vanzella et al. 2010; Vasei et al. 2016; Grazian et al. 2017; Japelj et al. 2017) of high-redshift galaxies. Here we are interested in constraining the second quantity, $\xi_{\rm ion}$. The primary way of determining $\xi_{\rm ion}$ is to infer the ionizing UV flux from the hydrogen recombination lines (e.g., H$\alpha$ or H$\beta$) assuming that the interstellar medium (ISM) is optically thick to ionizing photons and does not allow them to escape the galaxy. In this case, the rate of ionizations and, thus, the ionizing photon production rate can be inferred from recombination lines assuming case B recombination. As such, Bouwens et al. (2016), Nakajima et al. (2016), Matthee et al. (2017), Shivaei et al. (2018), and Tang et al. (2019) evaluated

* Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. This research is based on observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with programs 12201, 12931, 13389, and 14209.

$\xi_{\rm ion}$ as the ratio of hydrogen recombination lines to 1500 Å UV fluxes. Another indirect way of inferring $\xi_{\rm ion}$ is to implement metal nebular emission lines and stellar continuum into the photoionization models and output the shape of the ionizing spectrum and, thus, the best $\xi_{\rm ion}$ match to the observed spectrum (Stark et al. 2015, 2017; Chevallard et al. 2018).

However, all of these studies measure $\xi_{\rm ion}$ of high-redshift galaxies that are exclusively luminous H$\alpha$ or Ly$\alpha$ emitters or have extreme optical nebular emission lines. As such, there are not many measurements of $\xi_{\rm ion}$ in low-luminosity, low-mass galaxies (Lam et al. 2019).

It is not clear what type of galaxies contribute the most to the total ionizing photon budget necessary for reionization. Some studies suggest that perhaps rare Lyman continuum leakers with substantial star formation surface densities have led to a rapid, recent reionization at $z \sim 6$ (Naidu et al. 2020). Other studies predict that low-mass galaxies should have a greater contribution to reionization because of the steep faint-end slope of the $\rm UV_{1500}$ luminosity function of high-redshift galaxies (Reddy & Steidel 2009; Bouwens et al. 2012; Alavi et al. 2014; Finkelstein et al. 2015; Ishigaki et al. 2015; Atek et al. 2015; Livermore et al. 2017; Mehta et al. 2017). Additionally, at low mass, more ionizing photons are thought to escape from the galaxies into the IGM (Paardekooper et al. 2013; Wise et al. 2014; Erb 2015; Henry et al. 2015; Anderson et al. 2017; Karman et al. 2017) at high redshifts, possibly through hot "chimneys" created by feedback-driven outflows. In order to determine whether low-mass galaxies are the primary reionizing agents, we still need to investigate the ionizing photon production efficiency ($\xi_{\rm ion}$) of these low-mass galaxies and compare to their massive counterparts. However, despite its great importance, little is known about the $\xi_{\rm ion}$ in faint low-mass systems.

In this paper we measure, for the first time, $\xi_{\rm ion}$ for low-mass ($7.8 \leqslant \log(M_*) < 9.8$), low-luminosity ($-22 < M_{\rm UV} < -17.3$) galaxies at $1.4 < z < 2.7$. These galaxies may be intermediate-redshift analogs of the sources of reionization at $z > 6$. Galaxies in our sample are highly magnified by gravitational lensing by foreground galaxy clusters. The magnification enables us to detect low-luminosity galaxies, up to an intrinsic UV magnitude of $-17$. We quantify $\xi_{\rm ion}$ using H$\alpha$ recombination emission and nonionizing (1500 Å) UV fluxes from deep Keck/MOSFIRE spectroscopy and Hubble Space Telescope (HST) imaging, respectively. We also have H$\beta$ detections for all of the galaxies in our sample, which allows us to correct H$\alpha$ fluxes for the dust extinction via the Balmer decrement. We correct the UV stellar continuum using the dust extinction inferred from the spectral energy distribution (SED) fitting. We carefully select galaxies to be complete in both low and high UV luminosities.

There is an intrinsic scatter in the ratio of H$\alpha$ (or H$\beta$) to UV, especially in low-mass galaxies (Lee et al. 2009; Weisz et al. 2012; Domínguez et al. 2015; Guo et al. 2016; Emami et al. 2019). Many factors are known to contribute to this scatter, including bursty star formation, galaxy-to-galaxy dust extinction variation, escape of ionizing photons, varying initial stellar mass function (IMF), different stellar metallicities, and stellar models (Iglesias-Páramo et al. 2004; Boselli et al. 2009; Lee et al. 2009; Meurer et al. 2009; Weisz et al. 2012; Guo et al. 2016; Emami et al. 2019). As a result, we expect to see a similar scatter in the $\xi_{\rm ion}$ distribution, which makes it crucial to come up with an appropriate way of combining the galaxies' fluxes and derive a single $\xi_{\rm ion}$ value that properly represents the entire sample. Here we also address this issue and introduce a new way of stacking H$\alpha$ and UV fluxes that deals with the $\xi_{\rm ion}$ scatter in low-mass galaxies.

Since $\xi_{\rm ion}$ is related to the ionizing radiation intensity of the galaxies, it can also be inferred from other physical quantities that are also dependent on the ionizing radiation intensity, such as UV spectral slope (Robertson et al. 2013; Bouwens et al. 2015b; Duncan & Conselice 2015) and equivalent widths of nebular UV and optical emission lines (Stark et al. 2015, 2017; Chevallard et al. 2018; Tang et al. 2019). We investigate the relationship between $\xi_{\rm ion}$ and these physical quantities in our sample and see if the relations shown by previous works further extend to lower luminosity.

The outline of the paper is as follows. We describe the sample selection and data acquisition in Section 2. In Section 3 we present flux measurements. In Section 4 we describe two approaches of stacking fluxes and discuss the relevance of each for the $\xi_{\rm ion}$ determination. In Section 5 we show our results and compare them with previous works. We provide physical interpretations explaining our results in conjunction with previous studies in Section 6. Lastly, we conclude with a brief summary in Section 7. We assume a $\Lambda$-dominated flat universe with $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$, and $H_0 = 70\,{\rm km\,s^{-1}}$ ${\rm Mpc^{-1}}$. All magnitudes in this paper are in the AB system (Oke & Gunn 1983), and all equivalent widths are quoted in the rest frame.

## 2. Data

### 2.1. HST Data

Our sample is drawn from an HST survey (Alavi et al. 2016) that identifies faint star-forming galaxies at $1 < z < 3$ behind three lensing clusters—A1689 and two Hubble Frontier Fields (HFF) clusters, MACS J0717 and MACS J1149 (Lotz et al. 2017). The data reduction and photometric measurements are discussed in detail in Alavi et al. (2016). For galaxies behind A1689, we measure flux in eight photometric bands spanning the observed near-UV and optical. For galaxies behind MACS J0717 and MACS J1149 we measure flux in nine photometric bands spanning the observed near-UV, optical, and near-IR. The near-UV data (program IDs 12201, 12931, 13389) allow for efficient identification of the Lyman break, enabling accurate photometric redshifts at $1 < z < 3$.

We require a lens model for each cluster to correct for the lensing magnification and derive the intrinsic galaxy properties. As discussed in Alavi et al. (2016), for A1689 we use the lens model of Limousin et al. (2007), and for the HFF clusters we use the released models from the CATS[6] team (Jauzac et al. 2016 and Limousin et al. 2016 for MACS J1149 and MACS J0717, respectively). According to Priewe et al. (2017), for the HFF clusters and a typical magnification of our galaxies, which is around 5, the systematic error in the estimated magnification of different lens models is ~40%, which is small compared to our $M_*$ or UV luminosity bin sizes, which are about one order of magnitude (Figures 3 and 4). In Alavi et al. (2016) there is a full description of the lens models used for the A1689 and HFF clusters.

### 2.2. Spectroscopic Sample and Data Reduction

We obtain the rest-frame optical spectra of our sample via Keck/MOSFIRE observation. We select our spectroscopic

[6] https://archive.stsci.edu/prepds/frontier/lensmodels/

sample such that the bright rest-frame optical nebular emission lines fall within the atmospheric windows at $1.37 < z < 1.70$ and $2.09 < z < 2.61$. When selecting targets, we prioritized galaxies with high magnification and brighter observed optical flux densities ($M_B < 26.5$). The data were collected between 2014 January and 2017 March. Masks were made for the $1.37 < z < 1.70$ and $2.09 < z < 2.61$ redshift ranges, and all of the strong optical emission lines (H$\alpha$, [N II], [O III], H$\beta$, and [O II]) were targeted. For the lower-redshift mask, *Y*-, *J*-, and *H*-band spectroscopy was obtained. For the higher-redshift mask, *J*-, *H*-, and *K*-band spectroscopy was obtained. The total exposure times for each mask and filter range from 48 to 120 minutes. The typical FWHM seeing of our MOSFIRE spectra in any given mask and filter is $\sim 0\farcs71$. The slit widths are also $0\farcs70$.

The MOSFIRE data were reduced using the MOSFIRE Data Reduction Pipeline[7] (DRP). The DRP produces a 2D flat-fielded, sky-subtracted, wavelength-calibrated, and rectified spectrum for each slit. It also combines the spectra taken at each nod position (we used an ABBA dither pattern). The wavelength calibration for the *J*- and *H*-band spectra was performed using the skylines and for the *Y*- and *K*-band spectra a combination of skylines and neon lines. We then utilize custom IDL software, BMEP,[8] from Freeman et al. (2019) for the 1D extraction of spectra. The flux calibration is done in two stages. First, we use a standard star with spectral type ranging from B9 V to A2 V, which has been observed at similar airmass as the mask, to determine a wavelength-dependent flux calibration. We then do an absolute flux calibration using the spectrum of a star to which we assigned a slit in each mask.

## 3. Measurements

### 3.1. SED Fitting

Stellar masses, star formation rates, and stellar dust attenuation for our galaxies are estimated with SED fits to the photometry. Specifically, for the A1689 cluster, we use eight broadband filters spanning the observed near-UV to optical in the F225W, F275W, F336W, F475W, F625W, F775W, F814W, and F850LP filters. In addition, we use the photometry in two near-IR HST bands (F125W and F160W), though the imaging does not cover the full area covered by the near-UV and optical imaging.

For the two HFF clusters, we fit to nine broadband filters spanning the observed near-UV to near-IR in the F275W, F336W, F435W, F606W, F814W, F105W, F125W, F1140W, and F160W filters.

We use the stellar population fitting code FAST (Kriek et al. 2009), with the BC03 (Bruzual & Charlot 2003) population synthesis models and assume an exponentially increasing star formation history (which has been shown to best reproduce the observed star formation rates at $z \sim 2$; Reddy et al. 2012) with a Chabrier IMF (Chabrier 2003). As suggested by Reddy et al. (2018) for high-redshift low-mass galaxies, we use the SMC dust extinction curve (Gordon et al. 2003) with $A_V$ values varying between 0.0 and 3.0. We leave the metallicity as a free parameter between [0.4–0.8] $Z_\odot$. The age and star formation timescales can vary between $7.0 < \log(t)[\mathrm{yr}] < 10$ and $8.0 < \log(\tau)[\mathrm{yr}] < 11.0$, respectively. The redshifts are fixed to the values obtained spectroscopically. The $1\sigma$ confidence intervals are derived from a Monte Carlo method of perturbing the broadband photometry within the corresponding photometric uncertainties and refitting the SED 300 times. We note that we correct the broadband photometry for the contamination from the nebular emission lines using the line fluxes measured from the MOSFIRE spectra.

### 3.2. Emission-line Fitting

Spectral fitting was performed in each filter covering a galaxy, and for all of the aforementioned strong rest-optical emission lines, using the Markov Chain Monte Carlo Ensemble sampler, `emcee`[9] (Foreman-Mackey et al. 2013). Before fitting, to account for skyline contamination within a given spectrum, we removed any data points that have a corresponding error $>3\times$ the median error of the spectrum. Emission lines within close proximity to each other in a filter (e.g., [O III] $\lambda\lambda$4959, 5007 and H$\beta$; [N II] $\lambda\lambda$6548, 6583 and H$\alpha$) were fit simultaneously with single-Gaussian profiles, and the continuum was fit with a line. For the emission lines relevant to this paper (H$\beta$, [O III] $\lambda$5007, and H$\alpha$), the free parameters of the fits comprised the slope and y-intercept of the continuum line, a single emission-line width ($\sigma$) and redshift for the filter, and the amplitudes ($A_\lambda$) of the individual lines. When fitting the portion of the spectrum containing H$\beta$ and the [O III] doublet, the amplitude of [O III] $\lambda$5007 was set as a free parameter with the amplitude of [O III] $\lambda$4959 constrained to follow the intrinsic flux ratio of the doublet's lines: [O III] $\lambda5007/$[O III] $\lambda4959 = 2.98$ (Storey & Zeippen 2000). In the instances where [O III] $\lambda$5007 fell outside our spectroscopic coverage, its flux was determined with this flux ratio and the [O III] $\lambda$4959 line. The final spectroscopic redshift of a galaxy was determined via the weighted average of the redshifts fit to the different filters. More details about the spectroscopic line measurements of [O II] and other, fainter optical lines not considered here can be found in Gburek et al. (2019). To assess the quality of the fits to the spectra, posterior histograms were output for each free parameter (as well as for the fluxes), and 68% confidence intervals were fit to the histograms.

### 3.3. Slit-loss Correction

The emission-line fluxes need to be corrected for slit losses. This procedure is more important for extended or stretched (highly magnified) objects, as the slit may not fully cover the object. This needs to be done for each object in each MOSFIRE band and each mask. We adopt the following procedures: (1) we cut a $30'' \times 30''$ postage stamp centered on the galaxy from the F625W, as this filter gives a high signal-to-noise ratio image of the rest-frame UV light, and therefore the approximate morphology of the star-forming regions. (2) We identify the pixels corresponding to the object using the segmentation map output by Source Extractor (Bertin & Arnouts 1996). (3) We mask out all pixels of the nearby objects and background from the postage stamp and replace them with zero flux. (4) The sum of the total flux from pixels belonging to the object gives us the actual flux that Source Extractor measured. (5) We smooth the postage stamp applying a Gaussian kernel with an

[7] https://keck-datareductionpipelines.github.io/MosfireDRP/
[8] https://github.com/billfreeman44/bmep
[9] https://emcee.readthedocs.io/en/v2.2.1/

FWHM that is given by

$$\mathrm{FWHM}^2_{\mathrm{kernel}} = \mathrm{FWHM}^2_{\mathrm{seeing}} - \mathrm{FWHM}^2_{\mathrm{F625W}}. \quad (2)$$

$\mathrm{FWHM}^2_{\mathrm{seeing}}$ is the FWHM of the Gaussian fit to the profile of the slit star in the corresponding mask and filter. $\mathrm{FWHM}^2_{\mathrm{F625W}}$ is the FWHM of the F625W point-spread function (0.″1). This artificially degrades the resolution of the HST image to the same spatial resolution as the MOSFIRE observation. (6) We overlay the slit on the postage stamp of the smoothed image using its position angle, center, length, and width and block out regions of the object that fall out of the slit. (7) We sum the flux of the remaining pixels in the slit and denote it as in-slit flux. (8) We then determine a multiplicative factor required to have the in-slit flux match the total flux. This factor is the slit-loss correction and is applied to all lines in the corresponding filter and mask.

### 3.4. Sample Selections

There are 62 galaxies in our sample for which we have sufficient HST filter coverage spanning the observed near-UV to near-IR that enables a robust SED fit and, thus, reliable estimates of stellar properties (stellar mass, $V$-band dust attenuation ($A_V$), UV spectral slope ($\beta$), etc.). We remove some galaxies from the sample for the following reasons.

Noncovered H$\alpha$ or large H$\alpha$ errors: We remove 16 galaxies that do not have good H$\alpha$ measurements, because either H$\alpha$ is out of the wavelength coverage or the errors are significantly larger than the H$\alpha$ error distribution of the sample ($>10^{-17}$ erg s$^{-1}$ cm$^{-2}$, corresponding to a typical error $>1.76\times$ the median error, typically due to strong skyline contamination). The median H$\alpha$ flux error of the sample is $6.08 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$. We choose to impose a flux error cutoff rather than a signal-to-noise cutoff on our H$\alpha$ measurement, in order to avoid a bias against intrinsically faint H$\alpha$ emitters.

Galaxies with very high magnification: If a galaxy has a high average magnification, it means that it is sitting close to the caustic in the source plane. Thus, the gradient of the magnification can be large, resulting in large magnification differences across the galaxy. This could result in an observed ratio of $L_{\mathrm{H}\alpha}$ to $L_{\mathrm{UV}}$ that is different from the true ratio. Not only would this increase the scatter, but it can result in a bias, as the galaxies are selected via rest-frame UV continuum luminosity density. Hence, we remove seven galaxies whose magnifications ($\mu$) are $\mu > 30$ in A1689 and $\mu > 15$ in HFF clusters. The reason for choosing a larger magnification cut for the A1689 compared to other HFF clusters is because A1689 has a large Einstein radius, which provides high magnification over a large area in the source plane. Therefore, objects with a high magnification in A1689 are not required to be close to the critical lines, where the magnification formally diverges (Alavi et al. 2016).

Multiply imaged galaxies: We remove multiple images of two galaxies to avoid double-counting. In these cases, we keep the most highly magnified image in the sample unless the magnification is very large ($>30$ in A1689 and $>15$ in HFF clusters), in which case we use the next-brightest image. These multiple images were identified using Lenstool (Alavi et al. 2016; Limousin et al. 2016).

High slit-loss galaxies: For larger, extended galaxies, the slit-loss correction can be large, and the MOSFIRE measurement will only be sampling a small, possibly unrepresentative portion of the whole galaxy.

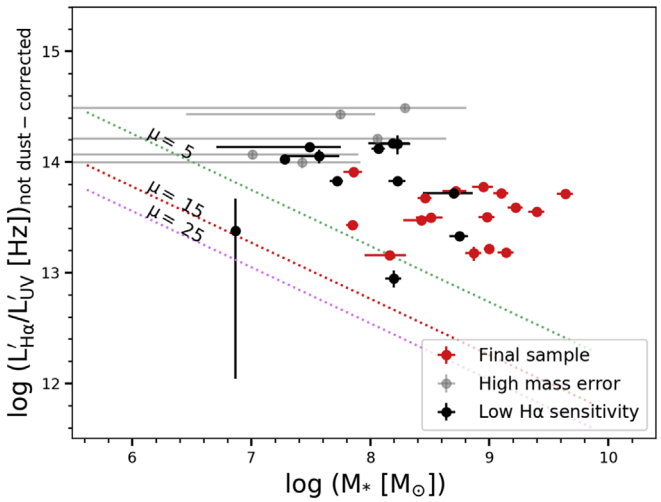


**Figure 1.** Non-dust-corrected $\log(L'_{\mathrm{H}\alpha}/L'_{\mathrm{UV}})$ as a function of $\log(M_*)$ derived from the observed $L_{\mathrm{H}\alpha}$ and $L_{\mathrm{UV}}$. The gray points show galaxies with high mass errors. Black points indicate galaxies that could not be detected if they had the very low observed $\log(L'_{\mathrm{H}\alpha}/L'_{\mathrm{UV}}) < 13.2$. The green, red, and magenta diagonal dotted lines indicate the typical $\log(L'_{\mathrm{H}\alpha}/L'_{\mathrm{UV}})$ detection limit for three magnification factors of 5, 15, and 25, respectively, below which galaxies are intrinsically too faint to be detected through MOSFIRE. The remaining galaxies in red are free of any biases in our measurements.

As such, we remove four galaxies with H$\alpha$ slit losses $>70\%$ from the sample. It is also worth noting that the typical slit loss of our sample is 40%.

Galaxies with large mass errors: We also make sure not to include galaxies with large stellar mass errors in our analysis. There are only four galaxies that lack HST rest-frame near-IR filter coverage and ultimately end up having large mass errors, shown as gray points in Figure 1. We note that we only exclude these galaxies from our sample when we perform flux stacking based on stellar masses (Section 5.1), but we use them when stacking based on properties other than the stellar mass (UV magnitude and UV spectral slope; Sections 5.2 and 5.3).

The final sample contains 28 galaxies that are free of the aforementioned concerns.

### 3.5. Non-dust-corrected $\xi_{\mathrm{ion}}$

The goal of this paper is to measure the ionizing photon production efficiency of galaxies ($\xi_{\mathrm{ion}}$) for our sample, which represents the rate of Lyman continuum photons per unit $\mathrm{UV}_{1500}$ luminosity as

$$\xi_{\mathrm{ion}} = \frac{Q_{\mathrm{H}^0}}{L_{\mathrm{UV}}}\ [\mathrm{s}^{-1}/\mathrm{erg\ s}^{-1}\,\mathrm{Hz}^{-1}], \quad (3)$$

where $L_{\mathrm{UV}}$ is the intrinsic UV continuum luminosity density (per unit frequency) around 1500 Å. Based on case B recombination, the rate of production of ionizing photons ($Q^0_{\mathrm{H}}$) can be determined from the hydrogen recombination lines, in this case H$\alpha$, as

$$L_{\mathrm{H}\alpha}[\mathrm{erg\ s}^{-1}] = 1.36 \times 10^{-12}\ Q_{\mathrm{H}^0}[\mathrm{s}^{-1}], \quad (4)$$

where $L_{\mathrm{H}\alpha}$ is the H$\alpha$ luminosity (Leitherer & Heckman 1995). Here we assume that all ionizing photons result in a photoionization (none escape into the IGM) and are converted

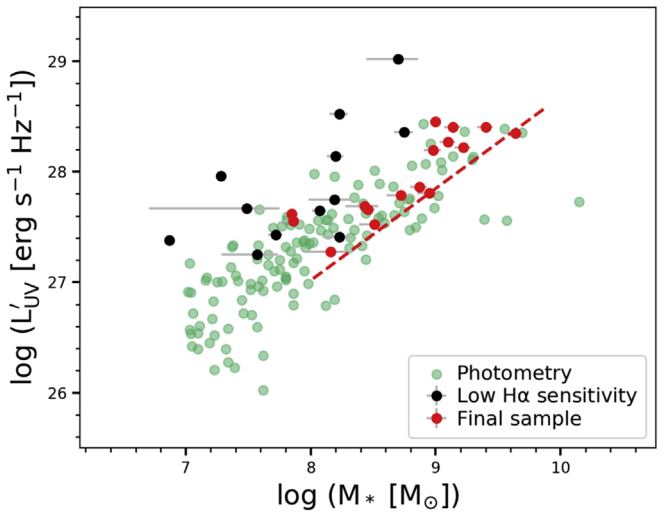


**Figure 2.** Non-dust-corrected log ($L'_{UV}$)–log($M_*$) relation of our lensed galaxies. Green points are the parent photometric sample with $M_*$ above $10^7 M_\odot$. The final spectroscopic $\xi'_{ion}$ sample is shown in red. Black points are removed from the $\xi'_{ion}$ sample owing to biases discussed in Section 3.4. The galaxies in our final sample (red points) span a similar range in $L'_{UV}$ at a given mass to the parent population, indicating that this sample is representative of low-mass galaxies at $1 < z < 3$. The red line denotes the lower edge of the log ($L'_{UV}$)–log($M_*$) main-sequence trend of the parent sample that is used to exclude galaxies with insufficient sensitivity (black points) from the $\xi'_{ion}$ sample described in Section 3.5.

into case B recombination emission. Therefore, the $\xi_{ion}$ values reported in this study are upper limits.

Figure 1 shows the ratio of observed (non-dust-corrected) $L'_{H\alpha}$ to $L'_{UV}$ as a function of stellar mass. (The prime sign on the $L'_{H\alpha}$, $L'_{UV}$, and $\xi'_{ion}$ is to distinguish them as the non-dust-corrected quantities.) Because the Hubble images are far more sensitive than our Keck/MOSFIRE observations, our primary incompleteness is determined by the depth of the spectroscopy. We are therefore concerned about completeness for galaxies with low $L'_{H\alpha}$ and, thus, low $\xi'_{ion}$. We therefore decide to only include galaxies in our final sample with spectra that are sensitive to the "worst-case" (lowest) observed $L'_{H\alpha}$ that can be expected.

In order to determine the lowest $L'_{H\alpha}$, we start by assuming the lowest $L'_{UV}$ for the measured stellar mass of the galaxy. This is found with a line near the lower edge of the observed log($L'_{UV}$)–log($M_*$) relation (at $M_* > 10^8$ $M_\odot$, where our sample is complete) shown in Figure 2. Once the worst-case $L'_{UV}$ is determined, we then assume a worst-case log($\xi_{ion}$) to find the faintest expected $L'_{H\alpha}$. We determine this worst-case value at the higher masses ($>10^9 M_\odot$, where we are complete), where we see that the lowest log($L'_{H\alpha}/L'_{UV}$) in our sample is $\sim$13.2. Finally, we compare this faintest $L'_{H\alpha}$ to our line sensitivity (assumed as $3\sigma$ H$\alpha$ flux detection) to determine what magnification is required to detect H$\alpha$ in our spectra. We keep all galaxies in our sample that have a high enough magnification. In this way, we ensure that all galaxies remaining in our sample have sufficiently sensitive spectra to detect galaxies with the lowest expected log($\xi_{ion}$).

Once we find the magnification threshold at any given mass, we remove galaxies in our sample whose magnifications are less than that threshold. There are 12 of these galaxies in our sample, which are shown as black points in Figure 1. Now we only work with the remaining objects (red points) in our sample, which are not affected by any biases. We note that log ($L'_{H\alpha}/L'_{UV}$) spans about 1 dex across the sample (13–14), as is evident in Figure 1.

We also perform a sanity check to determine whether our final sample can truly represent $\xi'_{ion}$ in low-mass galaxies or suffers from any biases against low-mass faint galaxies. This investigation is primarily due to the fact that our spectroscopic sample is a magnitude-limited subsample of our parent photometric sample ($B < 26.5$ AB). In this case, there is a possibility that we are populating the lower-mass bins only with the most luminous and youngest galaxies and might be missing the faint sources. To ensure that our final sample does not suffer from this bias, we plot the log ($L'_{UV}$)–log($M_*$) distribution of our parent photometric sample and compare it to the final $\xi'_{ion}$ sample in Figure 2. This figure indicates that our final sample has a similar distribution to the parent sample and is not biased toward high log($L'_{UV}$) values at a fixed stellar mass down to the mass of $10^{7.8} M_\odot$. Hence, our final $\xi'_{ion}$ sample is representative of low-mass galaxies at $1 < z < 3$ and is not biased against the low-mass, faint galaxies.

### 3.6. Dust Extinction Correction

We use the $A_V$ values derived from SED fits (Section 3.1) and assume an SMC extinction curve to correct for the dust attenuation of the UV luminosity density. We also use the Balmer decrement ($L'_{H\alpha}/L'_{H\beta}$) to determine the $L'_{H\alpha}$ attenuation assuming a Cardelli extinction curve (Cardelli et al. 1989).

## 4. Two Approaches to Flux Stacking for $\xi_{ion}$ Estimates

Here we attempt to evaluate the representative log($\xi_{ion}$) value of our sample. For this, we need to stack the dust-corrected H$\alpha$ and UV fluxes of individual galaxies. However, we note that the spread in log($\xi_{ion}$) is large ($\sim$1 dex). Given such a large spread in the *logarithm* of $\xi_{ion}$, we need to be careful about how we stack, depending on the question we are trying to answer.

There are two $\xi_{ion}$ values that we are interested in obtaining. First, we are interested in the properties of the *typical* galaxy, which can simply be obtained via the median, or the average of a symmetric distribution. Second, we are also interested in the total contribution of these galaxies to reionization, in which case we are interested in the *total* H$\alpha$ luminosity of all galaxies divided by the *total* UV luminosity of all galaxies. Such a number allows a direct conversion from UV luminosity functions to ionizing photon production rate densities. The stack in this case is *not* the average of the log($\xi_{ion}$) values that many have calculated before. Instead, this stack is equivalent to an $L_{UV}$-weighted average of the $L_{H\alpha}/L_{UV}$ ratios of the galaxies, as shown below.

$$\frac{\Sigma_i L_{H\alpha,i}}{\Sigma L_{UV,i}} = \frac{1}{\Sigma L_{UV,i}} \Sigma \frac{L_{UV,i} L_{H\alpha,i}}{L_{UV,i}}. \quad (5)$$

In order to obtain the composite log($\xi_{ion}$) for each of these methods more quantitatively, we follow the procedures below. For the first method we take the average of the logarithms of the ratio of $L_{H\alpha}$ to $L_{UV}$, and we refer to it as the *Standard* stacking method. For the second method we take the ratio of the average $L_{H\alpha}$ to the average $L_{UV}$ and then take the logarithm, and we refer to it as the *Effective* stacking method.

These two methods will give different $\xi_{ion}$ values for two reasons. First, since the *Standard* method takes the logarithm before averaging, it down-weights the importance of the

**Table 1**
log ($\xi_{\rm ion}$) Derived for Different Subsamples and Different Stacking Methods

| Subsample[a] | *Standard* log($\xi_{\rm ion}$)[b] | *Effective* log($\xi_{\rm ion}$)[c] | No. of Objects |
|---|---|---|---|
| $7.8 < \log(M_*) < 8.8$ | $25.17^{+0.13}_{-0.19}$ | $25.34^{+0.12}_{-0.15}$ | 6 |
| $8.8 < \log(M_*) < 9.8$ | $25.13^{+0.21}_{-0.19}$ | $25.39^{+0.14}_{-0.18}$ | 9 |
| $-22 < M_{\rm UV} < -19.5$ | $25.27^{+0.13}_{-0.16}$ | $25.47^{+0.10}_{-0.11}$ | 9 |
| $-19.5 < M_{\rm UV} < -17.3$ | $25.16^{+0.14}_{-0.18}$ | $25.47^{+0.12}_{-0.15}$ | 11 |
| $-2.4 < \beta < -1.75$ | $25.15^{+0.14}_{-0.23}$ | $25.46^{+0.11}_{-0.15}$ | 12 |
| $-1.75 < \beta < -0.93$ | $25.27^{+0.13}_{-0.17}$ | $25.45^{+0.13}_{-0.16}$ | 8 |

**Notes.**
[a] Log ($\xi_{\rm ion}$) measured for different subsamples of log($M_*$) (in units of $M_\odot$), UV magnitude, and UV continuum slope.
[b] log($\xi_{\rm ion}$) inferred from "*Standard*" stacking method.
[c] log($\xi_{\rm ion}$) inferred from "*Effective*" stacking method.

high-$\xi_{\rm ion}$ galaxies. Second, because the *Effective* method is effectively an $L_{\rm UV}$-weighted average of $\xi_{\rm ion}$, it may differ from the average if there is a correlation between $L_{\rm UV}$ and $\xi_{\rm ion}$ (see, e.g., Emami et al. 2019). The former method was used in Bouwens et al. (2016) and Shivaei et al. (2018), while the latter was used in Matthee et al. (2017) and Lam et al. (2019). It is therefore important to account for these different stacking methods when comparing to previous works.

We note here that our *Effective* method does not strictly get the true value of the total, volume-averaged log($\xi_{\rm ion}$) unless our sample galaxies also have similar luminosity and/or mass distributions to the true luminosity and/or mass functions. Of course, nearly all surveys of high-redshift galaxies have decreasing effective volumes at the faint end of the survey, but this is especially true for lensing surveys, which also rely on rarer, high magnifications at the faint end. Though this will remain a concern, we show in Section 5 that $\xi_{\rm ion}$ does not change significantly with luminosity or mass, so this additional uncertainty is likely to be small.

In order to get the uncertainties in the composite log($\xi_{\rm ion}$) of each stacking method, we use the bootstrap resampling technique: for a data sample of size $N$, we draw $N$ random values from the original sample and form a new sample of the same size and calculate its composite log($\xi_{\rm ion}$) the same way we did for the original sample. By repeating this 100,000 times, we build the distribution of the composite log($\xi_{\rm ion}$) values and calculate the 68% confidence interval of this distribution as the uncertainty in the composite log($\xi_{\rm ion}$). We also incorporate the errors in the H$\alpha$, H$\beta$, and UV fluxes in this calculation by drawing a random value from a normal distribution with a width equal to the $1\sigma$ error for each flux. In this way, we include the H$\alpha$ and H$\beta$ flux errors on the $A_{\rm H\alpha}$ determination and the H$\alpha$ and UV flux errors on the log ($\xi_{\rm ion}$) determination.

## 5. Results

In this section we discuss the relationship between $\xi_{\rm ion}$ and stellar mass, as well as other physical quantities such as UV magnitude, UV spectral slope ($\beta$), and the equivalent widths of nebular emission lines, and compare that with other studies. In Table 1, we present the composite log($\xi_{\rm ion}$) and its error in bins of stellar mass, UV magnitude ($M_{\rm UV}$), and UV continuum slope ($\beta$) obtained from the two stacking methods described in Section 4.

### 5.1. $\xi_{\rm ion}$ *versus Stellar Mass*

Galaxy stellar mass ($M_*$) is known to correlate with metallicity, which affects the stellar temperatures and, thus, $\xi_{\rm ion}$. We are therefore interested in examining the dependence of log($\xi_{\rm ion}$) on stellar mass for our sample. We present the log ($\xi_{\rm ion}$) derived from our two stacking methods as a function of log(stellar mass) in Figure 3.

As can be seen in Figure 3, the composite log($\xi_{\rm ion}$) is higher by at least 0.2 dex at all mass bins when using the *Effective* method compared to the *Standard* method. Since the errors in log($\xi_{\rm ion}$) of some galaxies are not negligible compared to the size of the spread in the log($\xi_{\rm ion}$) distribution of the sample, it might be thought that the stacked log($\xi_{\rm ion}$) derived from the *Effective* method is perhaps higher because of the large noise in these galaxies. We also check to make sure this enhancement is primarily due to the intrinsically high luminosities and not the noise. For that, we need to know how much the noise from our measurement has spread our observed log($\xi_{\rm ion}$) distribution. We run a simple simulation here: We first construct a normally distributed log($\xi_{\rm ion}$) of 1000 sources with an intrinsic spread of $\sigma_{\rm int}$ and perturb them with the fractional noise that is randomly drawn from the errors in $\xi_{\rm ion}$ of our sample. We then calculate the spread in this simulated $\xi_{\rm ion}$ distribution as $\sigma_{\rm sim}$. In order for the simulated spread to be equal to the observed spread (0.35), the intrinsic spread is required to be $\sim$0.29 dex, which implies $\sim$0.19 dex spread due to noise. Thus, the intrinsic spread is larger and is the primary reason for the increased log($\xi_{\rm ion}$) enhancement calculated via the *Effective* stacking method.

Now we compare our results with other studies at different redshifts or different stellar masses. First, we compare to a sample of local low-mass galaxies from Weisz et al. (2012). We have determined the composite log ($\xi_{\rm ion}$) of this sample in four mass bins, using the same two stacking methods we used for our sample, shown as green points in Figure 3. Similar to our sample, we see that the log($\xi_{\rm ion}$) measured from the *Effective* method is similar to or higher than the one derived from the *Standard* method in this sample. In particular, the difference between the two methods increases at lower masses, where the scatter in the log ($\xi_{\rm ion}$) is dramatic and is likely due to the increasing burstiness, as was found by Emami et al. (2019).

Comparing our results with Weisz et al. (2012), we find that at a given mass, our sample shows higher log($\xi_{\rm ion}$) relative to that of Weisz et al. (2012) (compare red markers with green ones), suggestive of a log($\xi_{\rm ion}$) evolution with redshift in the low-mass systems. We discuss possible explanations for this in Section 6.

We also compare our sample with higher-mass galaxies at similar redshift ($1.4 < z < 2.6$) from the MOSDEF Survey (Shivaei et al. 2018). In Figure 3 we show the log($\xi_{\rm ion}$) values for MOSDEF assuming SMC (Gordon et al. 2003) and Calzetti et al. (2000) UV dust extinction corrections. We see that the log ($\xi_{\rm ion}$) of our sample is in good agreement with that of Shivaei et al. (2018) at $10^9$–$10^{9.5} M_\odot$ within $1\sigma$ uncertainty, in the mass range where the two samples overlap. In fact, the log($\xi_{\rm ion}$) values of our galaxies in our sample and those at higher stellar mass are consistent at all stellar masses. Thus, there is no evidence for a trend in log($\xi_{\rm ion}$) with stellar mass from $10^{7.8}$–$10^{11} M_\odot$.

We also compare to the high-redshift sample of Lam et al. (2019), shown as purple circles in Figure 3. The sample is in the redshift range $3.8 < z < 5.3$. Galaxies in this sample are primarily selected to have Ly$\alpha$ emission in the MUSE data.

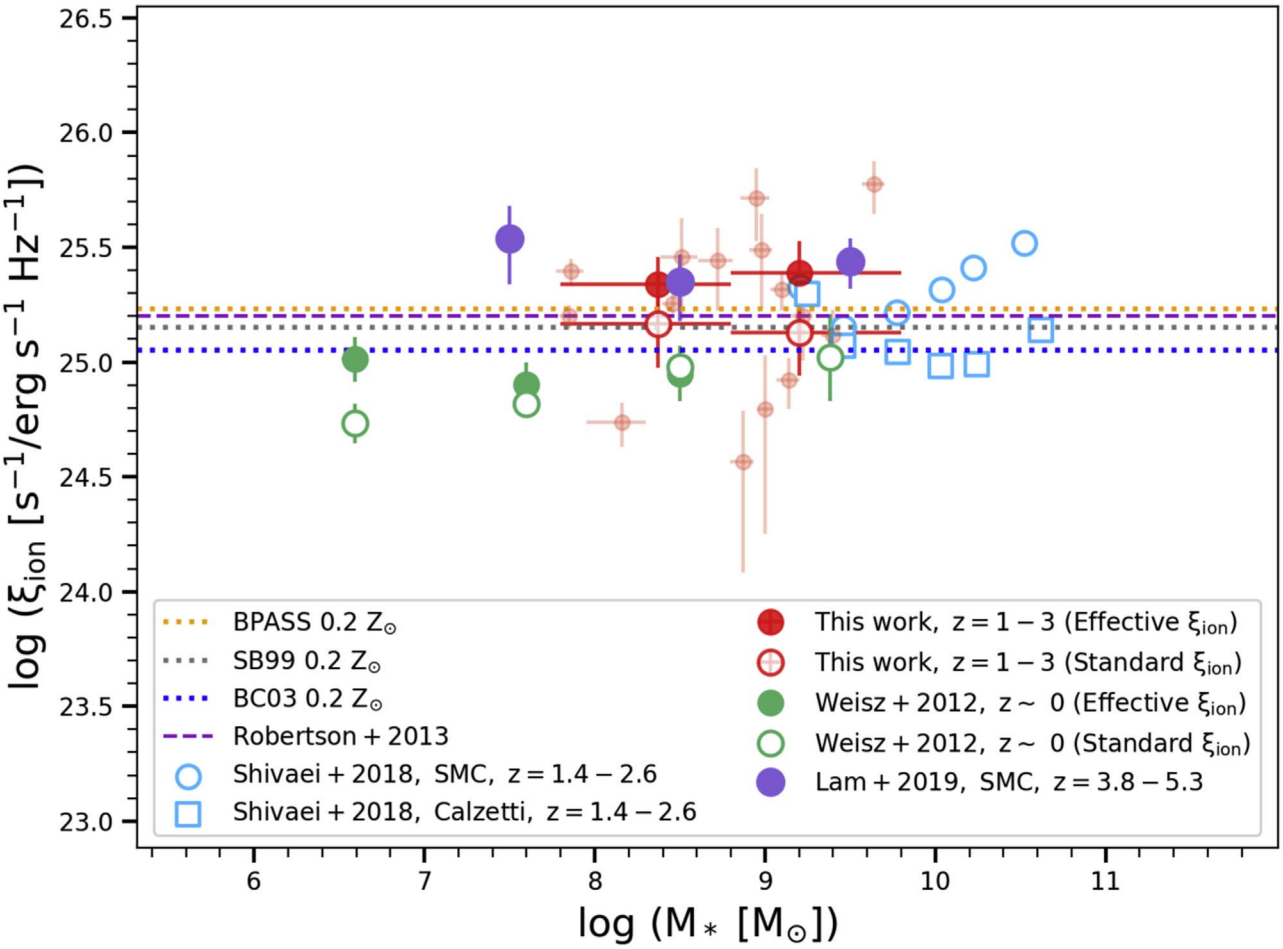


**Figure 3.** Log($\xi_{\rm ion}$) as a function of log($M_*$). Log($\xi_{\rm ion}$) derived from the *Standard* stacking method are shown with red open circles and the *Effective* stacking method with red filled circles. The log($\xi_{\rm ion}$) inferred from *Effective* method is ∼0.2 dex larger than that of the *Standard* method. Green open and filled circles denote the local sample of Weisz et al. (2012) applying the *Standard* and *Effective* stacking methods, respectively. Sky-blue open squares and circles denote the MOSDEF sample (Shivaei et al. 2018) of higher stellar mass galaxies using Calzetti et al. (2000) and SMC Gordon et al. (2003) UV dust corrections, respectively. Purple circles show the Lam et al. (2019) sample of faint ($L_{\rm UV} < 0.2\ L_*$) galaxies at higher redshifts ($z$ = 3.8–5.3). For a better comparison of samples with similar stacking methods, we use open markers to indicate the *Standard* stacking method and filled markers to indicate the *Effective* stacking method. The dashed line is the canonical value of 25.2 from Robertson et al. (2013). The local sample of Weisz et al. (2012) indicates lower log($\xi_{\rm ion}$) compared to ours. High-redshift samples of Shivaei et al. (2018) and Lam et al. (2019) lie within the 1$\sigma$ error bars of our two stacking methods. Orange, gray, and blue lines indicate the log ($\xi_{\rm ion}$) predicted by three different single stellar models, the BPASS model (Eldridge et al. 2017), Starburst99 (Leitherer et al. 2014), and BC03 (Bruzual & Charlot 2003), assuming a constant star formation history and 0.2$Z_\odot$ metallicity.

The sample includes galaxies of faint UV luminosities $-20.5 < M_{\rm UV} < -17.5$, similar to the galaxies in our intermediate-redshift sample. The log($\xi_{\rm ion}$) is inferred from the H$\alpha$ equivalent width, which in turn is derived from a power-law model spectrum fit through the flux of stacked Spitzer/IRAC [3.6]–[4.5] bands. The derived H$\alpha$ is then divided by the stacked UV fluxes. To that end, their way of log($\xi_{\rm ion}$) determination is similar to our *Effective* stacking method. The log($\xi_{\rm ion}$) obtained from the *Effective* method in our sample is consistent with that of Lam et al. (2019) within 1$\sigma$ error (compare red and purple filled markers).

### 5.2. $\xi_{\rm ion}$ versus UV Absolute Magnitude

$M_{\rm UV}$ is one of the easiest observables to obtain for high-redshift galaxies. Furthermore, the integral of the UV luminosity function is a critical calculation in determining the ionizing emissivities of galaxies. Therefore, we are particularly interested in whether or not there is any correlation between $M_{\rm UV}$ and $\xi_{\rm ion}$.

In Figure 4 we plot log($\xi_{\rm ion}$) as a function of $M_{\rm UV}$. We determine log($\xi_{\rm ion}$) for two bins of $M_{\rm UV}$ ($-22 < M_{\rm UV} < -19.5$ and $-19.5 \leqslant M_{\rm UV} < -17.2$) using the *Standard* and *Effective* stacking methods as were described in Section 4.

As in the previous section, we also need to take care that we only include galaxies for which we could detect very low $L_{\rm H\alpha}$. However, in this case, we are sampling galaxies based on their $M_{\rm UV}$, so we do not need to add a step of assuming an $M_{\rm UV}$–$M_*$ relation. Instead, we simply determine which galaxies could be detected if they had the very low observed log($L_{\rm H\alpha}/L_{\rm UV}$) $\gtrsim$ 13.2 and use each galaxy's measured $L_{\rm UV}$.

As was mentioned earlier in Section 3.4, we return galaxies with high stellar mass errors to our sample, as their masses are irrelevant in this analysis.

Log($\xi_{\rm ion}$) derived from the *Standard* method is similar in the two $M_{\rm UV}$ bins (25.17 and 25.28), while the *Effective* method gives a log($\xi_{\rm ion}$) of 25.47 for both bins. In both $M_{\rm UV}$ bins, the *Effective* method gives log($\xi_{\rm ion}$) values ∼0.2 dex larger than that of the *Standard* method. We also show results from Shivaei et al. (2018), Bouwens et al. (2016), and Lam et al. (2019) in Figure 4.

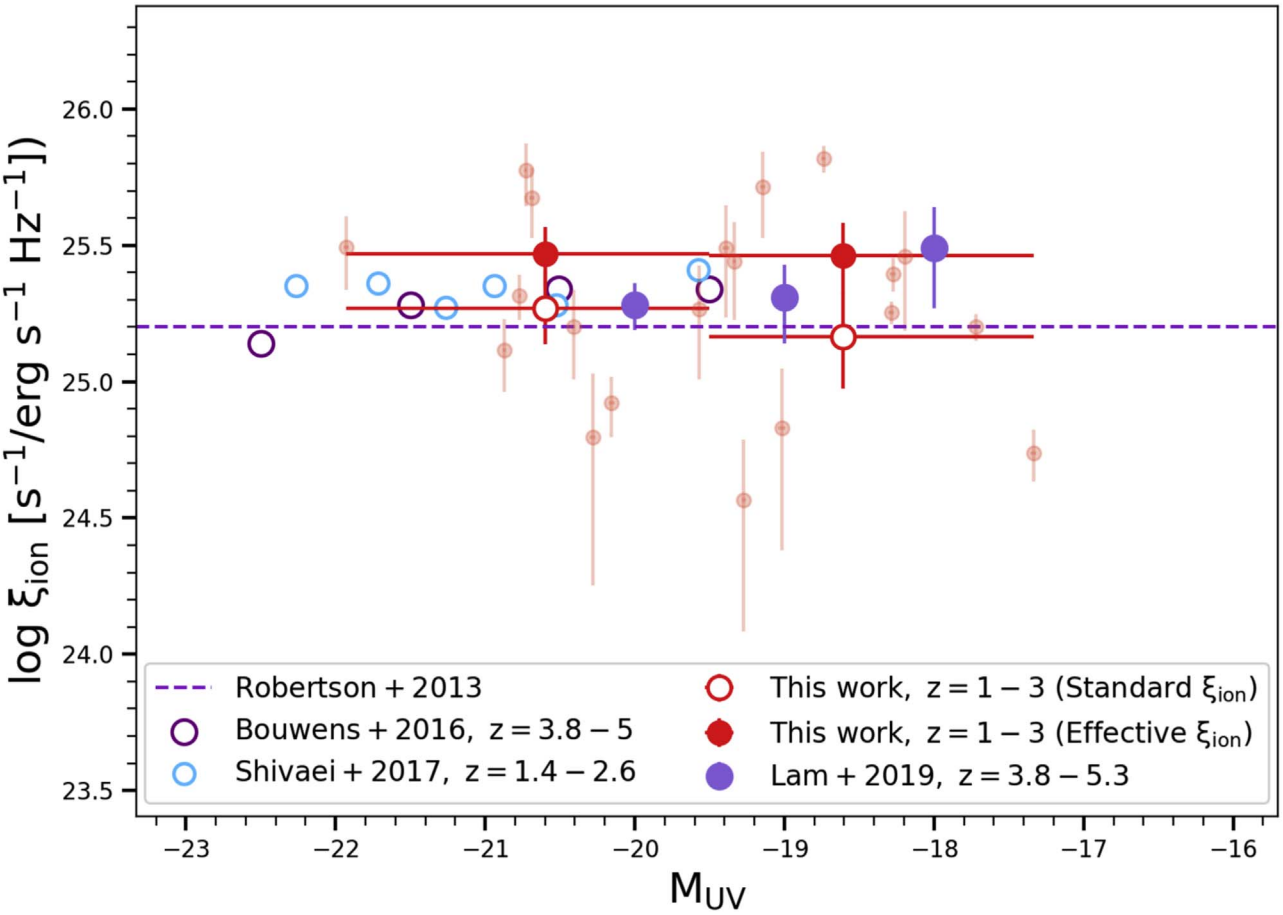


**Figure 4.** Log($\xi_{\rm ion}$) as a function of UV magnitude, $M_{\rm UV}$. Small, light-red circles denote individual galaxies in our sample. The large open and filled red circles show the log($\xi_{\rm ion}$) derived from the *Standard* and *Effective* stacking methods, respectively. Sky-blue circles show the stacks from Shivaei et al. (2018) for more massive $z \sim 2$ galaxies and an SMC UV dust correction. Dark- and light-purple circles denote Bouwens et al. (2016) and Lam et al. (2019) samples at $z \sim 4$–5, respectively. Similar to Figure 3, for a better comparison of samples with similar stacking methods, we use open markers to indicate the *Standard* stacking method and filled markers to show the *Effective* method. Our values agree with other studies within $1\sigma$ significance when the same stacking method as ours is used. No significant dependence of log($\xi_{\rm ion}$) with $M_{\rm UV}$ is found.

Comparing log($\xi_{\rm ion}$) of all works with analogous stacking techniques, we find that our values are in agreement with other works within $1\sigma$ significance.

We do not find any evidence of significant dependence of log($\xi_{\rm ion}$) on $M_{\rm UV}$ in our sample, in agreement with these other studies.

### 5.3. $\xi_{\rm ion}$ *versus UV Continuum Slope*

The UV continuum slope, $\beta$, is related to both the metallicity and age of the stellar populations and therefore, the inferred ionization capability of a galaxy driven by its young star populations. Therefore, we investigate whether $\xi_{\rm ion}$ is correlated with the more easily observable $\beta$.

The individual log($\xi_{\rm ion}$) values versus $\beta$ are plotted in Figure 5. We split the sample into two bins of $\beta$ ($-2.4 < \beta < -1.75$ and $-1.75 \leqslant \beta < -0.9$) and apply the same two stacking methods at each bin as we used for log($M_*$) and $M_{\rm UV}$. We find a similar log($\xi_{\rm ion}$) of 25.45 and 25.47 for the *Effective* method and a log($\xi_{\rm ion}$) range of 25.18−25.3 for the *Standard* method. We do not see any evidence for log($\xi_{\rm ion}$) being correlated with $\beta$ in our sample. Again, we find that our log ($\xi_{\rm ion}$) stack values are consistent with those of other studies at similar or higher redshifts.

### 5.4. $\xi_{\rm ion}$ *versus* EW$_{\rm [O\,III]\,\lambda5007}$, EW$_{\rm H\alpha}$

Finally, we investigate the relationship between log($\xi_{\rm ion}$) and the equivalent widths of optical nebular emission lines. One expects a positive correlation because the optical line equivalent widths are directly related to the luminosity-weighted age of the stellar populations, which itself affects $\xi_{\rm ion}$.

This relationship was first investigated by Chevallard et al. (2018), who found that log($\xi_{\rm ion}$) in galaxies with strong ionizing emissivities are scaled with the equivalent width of the combined [O III] $\lambda\lambda$4959, 5007 lines. They showed this for a sample of local star-forming galaxies with very high rest-frame equivalent widths ($560 < {\rm EW}_{\rm [O\,III]\,\lambda5007} < 2370$ Å). Tang et al. (2019) confirmed the existence of such a scaling relation for a sample of 227 low-mass ($10^7 < M_*/M_\odot < 10^{10}$), [O III] emitters with $225 < {\rm EW}_{\rm [O\,III]\,\lambda5007} < 2500$ Å at $1.3 < z < 2.4$, suggesting that higher equivalent width systems are more efficient ionizing agents. Given that, we aim to test this for our galaxies to see whether this relation further extends to lower equivalent width systems or not. We calculate the EW$_{\rm [O\,III]\,\lambda5007}$ by taking the ratio of the [O III] emission-line flux to the flux of the rest-frame $\lambda$5007 continuum from our HST near-IR fluxes, which have been corrected for emission-line contamination. We show log ($\xi_{\rm ion}$) versus log [O III] $\lambda$5007 equivalent width (EW$_{\rm [O\,III]\,\lambda5007}$)

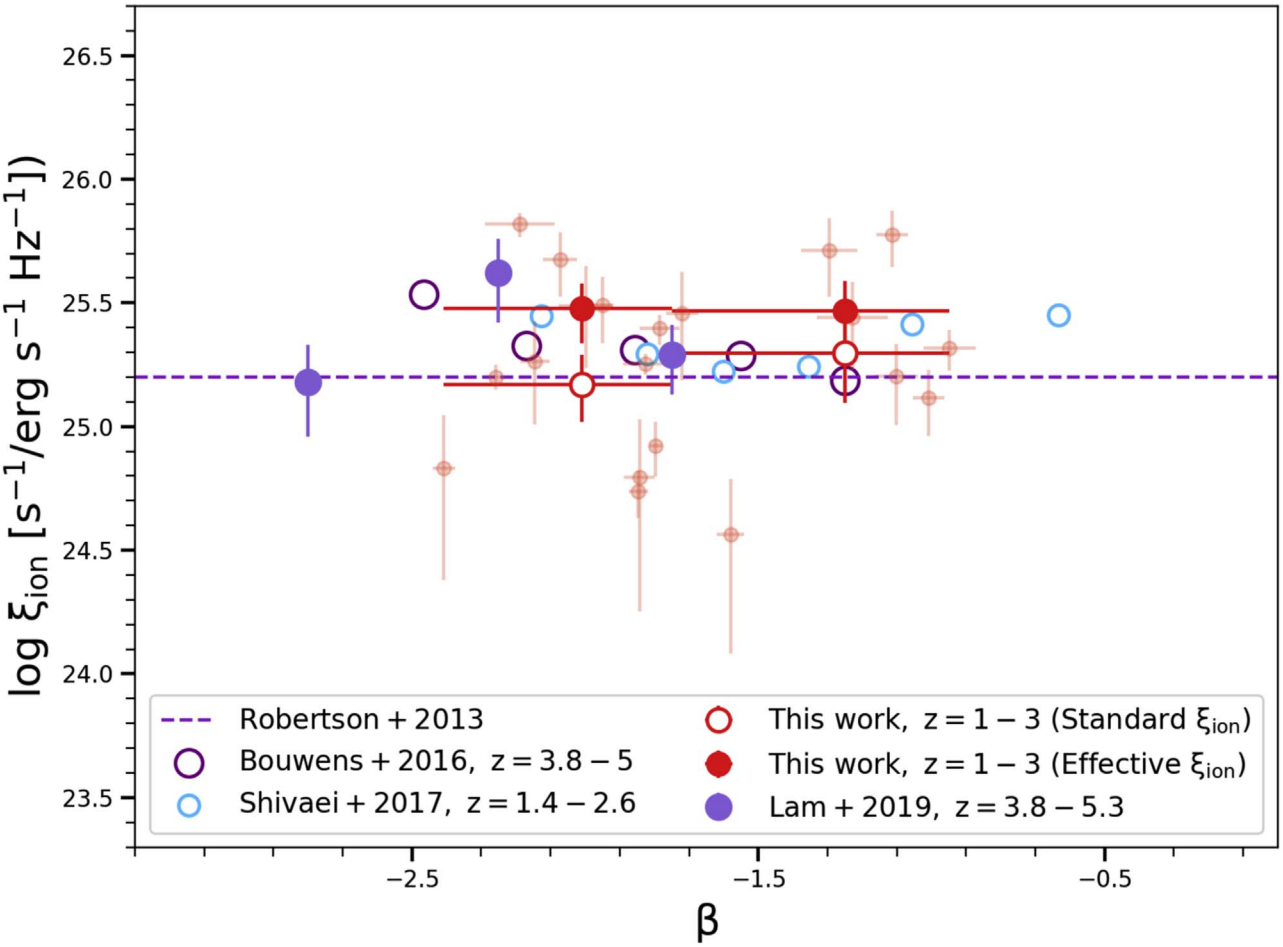


**Figure 5.** Log($\xi_{ion}$) as a function of UV slope $\beta$. The symbols are the same as in Figure 4. No dependence of log($\xi_{ion}$) with $\beta$ is seen in our sample.

in the top panel of Figure 6. Our galaxies span a large range of rest-frame equivalent widths ($20 < EW_{[O\,III]\lambda5007} < 1500$ Å) but generally extend lower than these previous studies. There is a trend of increasing log($\xi_{ion}$) with log($EW_{[O\,III]\lambda5007}$). To quantify this trend, we fit a line to the sample using ordinary least squares and plot the best fit, along with the 68% confidence region. We see a correlation between log($\xi_{ion}$) and log($EW_{[O\,III]\lambda5007}$) with a slope of $0.38 \pm 0.16$.

In addition, we overlay the trend from Tang et al. (2019) at larger $EW_{[O\,III]\lambda5007}$, which is steeper than ours, with smaller uncertainty in the fit. We note that our galaxies at $EW_{[O\,III]\lambda5007} > 200$ Å display a similar trend to that of Tang et al. (2019). The discrepancy between the two trends suggests that the slope in the log($\xi_{ion}$)–log($EW_{[O\,III]\lambda5007}$) gets shallower at lower equivalent widths.

We also plot the log($\xi_{ion}$) versus H$\alpha$ equivalent width ($EW_{H\alpha}$) relation in the bottom panel of Figure 6. After fitting a line through the points, we find a significant correlation, with a slope of $0.52 \pm 0.16$ between the two indicators. We further overplot the trend from Tang et al. (2019) and Faisst et al. (2019), which contains galaxies at $z \sim 4.5$ and stellar masses $>10^{9.7} M_{\odot}$. Our trend has a similar slope to those of Tang et al. (2019) and Faisst et al. (2019), but again with a larger uncertainty in the fit. Such a steep slope implies that log($\xi_{ion}$) is more correlated with $EW_{H\alpha}$ than with $EW_{[O\,III]\lambda5007}$, as was reported by Tang et al. (2019). In addition, Reddy et al. (2018) have also found similar trends of $EW_{H\alpha}$ and $EW_{[O\,III]\lambda5007}$ versus $\xi_{ion}$ to ours for more massive galaxies in the MOSDEF survey ($10^9 < M_*/M_{\odot} < 10^{10.5}$) at $1.4 < z < 3.8$. Tang et al. (2019) argue that the log($\xi_{ion}$)-$EW_{[O\,III]}$ and log($\xi_{ion}$)–$EW_{H\alpha}$ correlation should not hold at lower equivalent widths (below 200 Å). According to Tang et al. (2019), the equivalent widths correlate with $\xi_{ion}$ only within the first 100 Myr since the onset of star formation. After this time, both young (O-type) and intermediate-aged (B- and A-type) populations reach equilibrium, resulting in a constant $L_{H\alpha}$-to-$L_{UV}$ ratio and a plateau in $\xi_{ion}$ versus equivalent width. This is also evident in our sample as we get shallower slopes when including lower equivalent widths into the line fits (below 200 Å). However, this star formation history interpretation is only correct if one assumes a constant star formation history. A more comprehensive investigation of the $\xi_{ion}$ dependence on the equivalent widths requires additional analysis of the star formation histories of galaxies, as well as other physical properties, which is beyond the scope of this paper.

Finally, we note that this correlation between log($\xi_{ion}$) and the equivalent widths of some ionization-sensitive nebular emission lines can be used as a proxy for $\xi_{ion}$ at high redshifts when the direct measurement of rest-frame $L_{UV}$ is not available (Chevallard et al. 2018; Tang et al. 2019).

## 6. Discussion

In Section 5.1 we reported an increase in the log($\xi_{ion}$) of our $1.4 < z < 2.7$ sample relative to the low-redshift sample of Weisz et al. (2012) (see Figure 3). This suggests that at higher

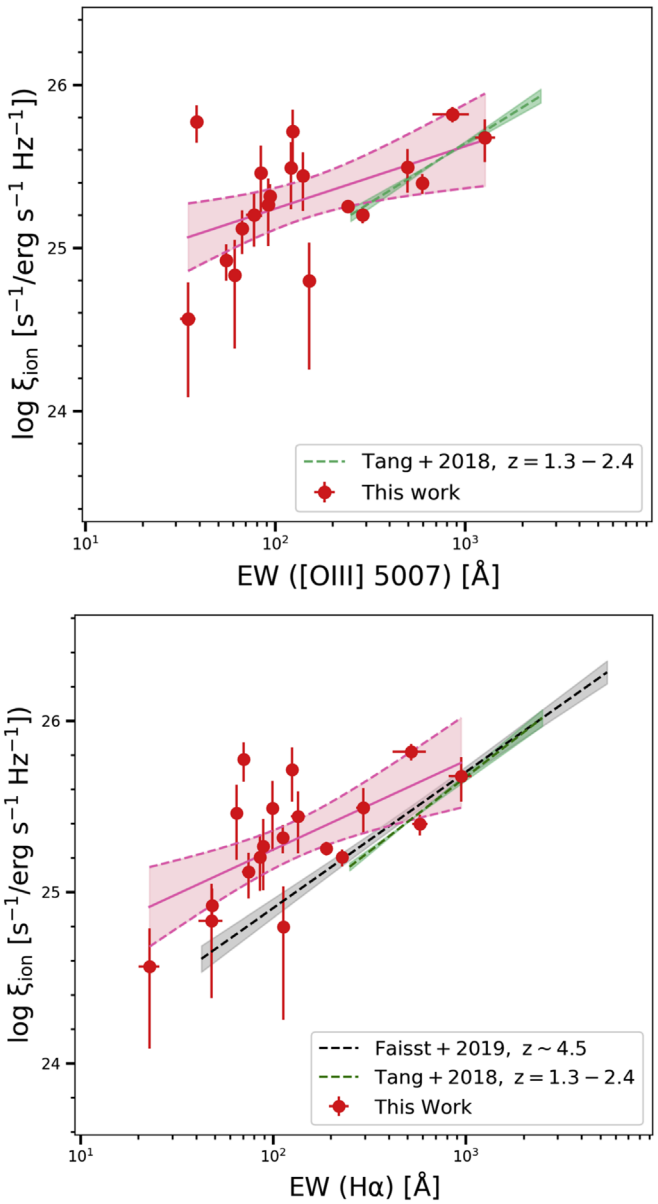


**Figure 6.** Top: log ($\xi_{\rm ion}$) vs. [O III] $\lambda$5007 equivalent width. The solid red line and the pink region denote the best-fit line and $1\sigma$ confidence region, respectively. The green dashed line is from Tang et al. (2019) for extreme [O III] emitters at $1.3 < z < 2.4$. Overall, a positive slope of $0.38 \pm 0.16$ is apparent between the two properties, but less steep than Tang et al. (2019). Bottom: log ($\xi_{\rm ion}$) vs. H$\alpha$ equivalent width ($\rm EW_{H\alpha}$). There is a slope of $0.52 \pm 0.16$ between the two properties. The gray line denotes the Faisst et al. (2019) relation at $z \sim 4.5$, which overlaps with the Tang et al. (2019) (green) but extends to a larger range of H$\alpha$ equivalent widths (40–5000 Å).

redshifts, galaxies with mass range of $10^{7.8} \leqslant M_{\odot} \leqslant 10^{9.5}$ produce more ionizing photons relative to the nonionizing UV photons when compared to their low-redshift counterparts. Here we provide possible explanations for this difference between high- and low-redshift samples.

First, the oxygen-to-iron abundance ratio of galaxies affects the production of ionizing photons at high redshift. Recent studies by Steidel et al. (2016) and Strom et al. (2018) show that in high-mass ($9 \leqslant log(M_*/M_{\odot}) \leqslant 10.8$), high-redshift galaxies ($z = 2.4 \pm 0.11$), the [O/Fe] abundance is supersolar ($\simeq$4–5 $[{\rm O/Fe}]_{\odot}$), referred to as "$\alpha$ enhancement." This has been shown in the composite UV spectrum of a representative sample of galaxies in the KBSS-MOSFIRE spectroscopic survey (Steidel et al. 2014). They found that emission spectra from photoionization modeling best match their composite UV spectra with stellar models with low stellar metallicities ($Z/Z_{\odot} \sim 0.1$), while the gas-phase oxygen abundances measured from nebular emission lines are $\sim$4 times higher. Given that stellar opacity is dominated by iron, this suggests a supersolar [O/Fe].

The deficit of iron in high-redshift galaxies can be explained by a model in which iron is produced during a delayed detonation of Type Ia supernovae (Khokhlov 1991). As a consequence, in high-redshift galaxies not all white dwarf stars have detonated and released iron into the ISM. Since iron predominantly controls the opacity of stellar atmospheres, its deficiency allows stars of a given mass to be hotter and, thus, have higher ionizing photon production at higher redshifts, leading to an increase in $\xi_{\rm ion}$ compared to local samples. It is likely the case that stellar populations of lower-mass galaxies at $z \sim 2$ are as young as the higher-mass galaxies of Steidel et al. (2014), and therefore exhibit a similar $\alpha$ enhancement. To confirm this requires measurement of the iron abundance of low-mass galaxies via absorption lines in their UV spectrum.

Second, this excess in the ionizing UV photons could be due to a recent increase in the star formation activity of high-redshift galaxies, resulting in an enhancement in the $L_{\rm H\alpha}$ relative to the $L_{\rm UV}$. This effect has also been reported in Faisst et al. (2019) such that their $z \sim 4.5$ main-sequence galaxies indicate a $\xi_{\rm ion}$ median of 25.5, which is 0.3 dex above the typically used canonical value of Robertson et al. (2013). This recent star formation activity can be in the form of continuous increase in the star formation history of the galaxy, which is typical in $z > 4$ galaxies (Behroozi et al. 2019), or a recent, rapid burst. Either of these star formation scenarios will lead to an increase in the number density of young stellar populations relative to the number density of intermediate-aged stellar populations in galaxies and ultimately results in an excess in the $L_{\rm H\alpha}$-to-$L_{\rm UV}$ ratios. However, exploring the effect of star formation variation on the ionizing photon production efficiency requires a deeper analysis of the star formation properties of galaxies at different epochs, which is beyond the scope of this paper and will be the subject of a future investigation.

We also investigate the $\xi_{\rm ion}$ predicted by different star formation synthesis models, BPASS (Eldridge et al. 2017), BC03 (Bruzual & Charlot 2003), and Starburst99 (Leitherer et al. 1999, 2014), for a constant star formation history (which should be equivalent to the average of many galaxies at various stages of burstiness) and compare them with our observed values as shown in Figure 3. Assuming a 0.2 $Z_{\odot}$ metallicity, Chabrier IMF (Chabrier 2003), and Padova isochrone (Bertelli et al. 1994; Bressan et al. 1993; Fagotto et al. 1994), we find that these models produce log($\xi_{\rm ion}$) values within 25–25.2. When including the effect of stripped, binary stars in the BPASS (Eldridge et al. 2017) and Starburst99 (Götberg et al. 2019) single-star models, we find only a small, $\sim$5%, enhancement to $\xi_{\rm ion}$. This is because these stripped, binary stars emit H I-ionizing photons at a rate that is 5% of the rate of H I-ionizing photons emitted by the massive single O-type stars (Götberg et al. 2019). As a result, when a constant star formation history is assumed, the emission from the massive single stars always dominates the emission from other less massive stripped, binary stars. Therefore, the evolution in $\xi_{\rm ion}$

cannot fully be explained by an evolution in binarity as predicted by these models.

## 7. Summary

In this paper we measure the ionizing photon production efficiency per unit 1500 Å UV luminosity, $\xi_{\rm ion}$, of a sample of low-mass ($7.8 < \log(M^*/M_\odot) < 9.8$) lensed galaxies at $1.4 < z < 2.7$. We obtained rest-frame optical spectroscopy of these faint sources that are magnified by the foreground lensing clusters A1689, MACS J0717, and MACS J1149, enabling us to extend the $\xi_{\rm ion}$ measurement to lower masses and fainter UV magnitudes ($M_{\rm UV} < -18$) than previously probed at these redshifts. We use the ratio of the H$\alpha$ luminosity (from Keck/MOSFIRE spectroscopy) and the 1500 Å UV luminosity density (from HST imaging) to measure $\xi_{\rm ion}$. We limit our sample to those objects where we are complete in our measurement of $\xi_{\rm ion}$.

We divide the sample into bins of different physical quantities such as stellar mass, absolute UV magnitude ($M_{\rm UV}$), and UV spectral slope ($\beta$) and calculate the stacked $\log(\xi_{\rm ion})$ in each bin using two different stacking methods. The most common method is to take the average of the log $(L_{\rm H\alpha}/L_{\rm UV})$ of galaxies to determine the standard $\log(\xi_{\rm ion})$ value, referred to as the "*Standard*" stacking method. The second method is to take the log of $\rm sum(L_{H\alpha})/sum(L_{UV})$, which we refer to as the "*Effective*" stacking method. This method is preferable when one is interested in calculating the total ionizing UV luminosity density from the nonionizing UV luminosity function. Here we list our main results:

1. In samples with a large spread in the $\log(\xi_{\rm ion})$ distribution, the stacked $\log(\xi_{\rm ion})$ from the two stacking methods can be significantly different. This is evident in the low-mass local sample of Weisz et al. (2012) in Figure 3.
2. We measure a value of $\log(\xi_{\rm ion})$ $\sim$25.47 $\pm$ 0.09 for our UV-complete sample in the range $-22 < M_{\rm UV} < -17.3$ and $\sim$25.37 $\pm$ 0.11 for our mass-complete sample in the range $7.8 < \log(M_*) < 9.8$. The slight difference between these two values is due to small differences in the samples.
3. We find that the $\log(\xi_{\rm ion})$ derived from the *Effective* method is about 0.2 dex higher than that of the *Standard* method in our sample of $z \sim 2$ galaxies, meaning that low UV luminosity systems may contribute $\sim$60% more ionizing photons than inferred from other stacking methods.
4. The measured $\log(\xi_{\rm ion})$ of our $z \sim 2$ sample is higher than the low-mass local sample of Weisz et al. (2012) by $\sim$0.2–0.3 dex when measured in a consistent manner. We argue that this can be attributed to different physical properties in high- and low-redshift galaxies: (i) delayed Type Ia supernovae results in an $\alpha$-enhancement (lower Fe relative to O) in the stellar population, which causes stars of a given mass to be hotter and, thus, have higher ionizing photon production (Steidel et al. 2016). (ii) An increase in the recent star formation activity of the high-redshift galaxies can also increase the relative number of young stars, thereby increasing the ratio of ionizing photons to nonionizing photons.
5. We find similar $\xi_{\rm ion}$ values to galaxies of higher mass at similar redshift (Shivaei et al. 2018) and similar mass at higher redshift (Lam et al. 2019). $\xi_{\rm ion}$ derived from these three samples are roughly consistent with the predictions of the single stellar models with an assumption of 0.2 $Z_\odot$ stellar metallicity.
6. We find no strong dependence between $\xi_{\rm ion}$ and $M_{\rm UV}$ or UV spectral slope, $\beta$, consistent with Bouwens et al. (2016), Shivaei et al. (2018), and Lam et al. (2019).
7. There is a positive correlation between $\xi_{\rm ion}$ and both H$\alpha$ and $[\rm O\,III]_{5007}$ equivalent widths in our faint, lower equivalent width systems. This confirms that the equivalent width of these strong optical lines can act as a proxy for $\xi_{\rm ion}$, though the relation appears to be less steep and with larger scatter at lower equivalent widths.
8. We find an intrinsic scatter of $\sim$0.29 dex in the $\log(\xi_{\rm ion})$ distribution of our sample. Many physical factors can cause this scatter. In a future paper we will investigate the underlying causes of this scatter in our lensed, high-redshift sample.

We thank the anonymous referee for providing useful comments that helped improve the quality of this paper. This material is based on work supported by the National Science Foundation under grant No. 1617013. Support for programs No. 12201, 12931, 13389, and 14209 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

D.R.W. acknowledges fellowship support from the Alfred P. Sloan Foundation and the Alexander von Humboldt Foundation.

*Facilities:* Keck:I (MOSFIRE), HST (WFC3, ACS) .

## ORCID iDs

Najmeh Emami https://orcid.org/0000-0003-2047-1689
Brian Siana https://orcid.org/0000-0002-4935-9511
Anahita Alavi https://orcid.org/0000-0002-8630-6435
Timothy Gburek https://orcid.org/0000-0002-7732-9205
William R. Freeman https://orcid.org/0000-0003-3559-5270
Johan Richard https://orcid.org/0000-0001-5492-1049
Daniel R. Weisz https://orcid.org/0000-0002-6442-6030

## References

Alavi, A., Siana, B., Richard, J., et al. 2014, ApJ, 780, 143
Alavi, A., Siana, B., Richard, J., et al. 2016, ApJ, 832, 56
Anderson, L., Governato, F., Karcher, M., Quinn, T., & Wadsley, J. 2017, MNRAS, 468, 4077
Atek, H., Richard, J., Kneib, J.-P., et al. 2015, ApJ, 800, 18
Becker, R. H., Fan, X., White, R. L., et al. 2001, AJ, 122, 2850
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&AS, 106, 275
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Boselli, A., Boissier, S., Cortese, L., et al. 2009, ApJ, 706, 1527
Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., & Franx, M. 2006, ApJ, 653, 53
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, ApJL, 752, L5

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015a, ApJ, 803, 34
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015b, ApJ, 811, 140
Bouwens, R. J., Smit, R., Labbé, I., et al. 2016, ApJ, 831, 176
Bressan, A., Fagotto, F., Bertelli, G., & Chiosi, C. 1993, A&AS, 100, 647
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chabrier, G. 2003, PASP, 115, 763
Chevallard, J., Charlot, S., Senchyna, P., et al. 2018, MNRAS, 479, 3264
Domínguez, A., Siana, B., Brooks, A. M., et al. 2015, MNRAS, 451, 839
Duncan, K., & Conselice, C. J. 2015, MNRAS, 451, 2030
Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, PASA, 34, e058
Emami, N., Siana, B., Weisz, D. R., et al. 2019, ApJ, 881, 71
Erb, D. K. 2015, Natur, 523, 169
Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994, A&AS, 105, 29
Faisst, A. L., Capak, P. L., Emami, N., Tacchella, S., & Larson, K. L. 2019, ApJ, 884, 133
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, AJ, 122, 2833
Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, ApJ, 879, 36
Finkelstein, S. L., Ryan, R. E., Jr., Papovich, C., et al. 2015, ApJ, 810, 71
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Freeman, W. R., Siana, B., Kriek, M., et al. 2019, ApJ, 873, 102
Gburek, T., Siana, B., Alavi, A., et al. 2019, ApJ, 887, 168
Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279
Götberg, Y., de Mink, S. E., Groh, J. H., Leitherer, C., & Norman, C. 2019, A&A, 629, A134
Grazian, A., Giallongo, E., Paris, D., et al. 2017, A&A, 602, A18
Guo, Y., Rafelski, M., Faber, S. M., et al. 2016, ApJ, 833, 37
Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, ApJ, 809, 19
Iglesias-Páramo, J., Boselli, A., Gavazzi, G., & Zaccardo, A. 2004, A&A, 421, 887
Inoue, A. K., Iwata, I., & Deharveng, J.-M. 2006, MNRAS, 371, L1
Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2015, ApJ, 799, 12
Japelj, J., Vanzella, E., Fontanot, F., et al. 2017, MNRAS, 468, 389
Jauzac, M., Richard, J., Limousin, M., et al. 2016, MNRAS, 457, 2029
Karman, W., Caputi, K. I., Caminha, G. B., et al. 2017, A&A, 599, A28
Khokhlov, A. M. 1991, A&A, 245, 114
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Lam, D., Bouwens, R. J., Labbé, I., et al. 2019, A&A, 627, A164
Lee, J. C., Gil de Paz, A., Tremonti, C., et al. 2009, ApJ, 706, 599
Leitherer, C., Ekström, S., Meynet, G., et al. 2014, ApJS, 212, 14
Leitherer, C., & Heckman, T. M. 1995, ApJS, 96, 9
Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3
Limousin, M., Richard, J., Jullo, E., et al. 2007, ApJ, 668, 643
Limousin, M., Richard, J., Jullo, E., et al. 2016, A&A, 588, A99
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97
Madau, P., Rees, M. J., Volonteri, M., Haardt, F., & Oh, S. P. 2004, ApJ, 604, 484
Mason, C. A., Naidu, R. P., Tacchella, S., & Leja, J. 2019, MNRAS, 489, 2669
Matthee, J., Sobral, D., Best, P., et al. 2017, MNRAS, 465, 3637
McGreer, I. D., Mesinger, A., & D'Odorico, V. 2015, MNRAS, 447, 499
Mehta, V., Scarlata, C., Rafelski, M., et al. 2017, ApJ, 838, 29
Meurer, G. R., Wong, O. I., Kim, J. H., et al. 2009, ApJ, 695, 765
Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, ApJ, 892, 109
Nakajima, K., Ellis, R. S., Iwata, I., et al. 2016, ApJL, 831, L9
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2013, ApJ, 773, 75
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Paardekooper, J. P., Khochfar, S., & Dalla, C. V. 2013, MNRAS, 429, L94
Priewe, J., Williams, L. L. R., Liesenborgs, J., Coe, D., & Rodney, S. A. 2017, MNRAS, 465, 1030
Reddy, N. A., Oesch, P. A., Bouwens, R. J., et al. 2018, ApJ, 853, 56
Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, ApJ, 754, 25
Reddy, N. A., & Steidel, C. C. 2009, ApJ, 692, 778
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19
Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71
Shivaei, I., Reddy, N. A., Siana, B., et al. 2018, ApJ, 855, 42
Siana, B., Teplitz, H. I., Colbert, J., et al. 2007, ApJ, 668, 62
Somerville, R. S., Bullock, J. S., & Livio, M. 2003, ApJ, 593, 616
Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469
Stark, D. P., Walth, G., Charlot, S., et al. 2015, MNRAS, 454, 1393
Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165
Steidel, C. C., Strom, A. L., Pettini, M., et al. 2016, ApJ, 826, 159
Storey, P. J., & Zeippen, C. J. 2000, MNRAS, 312, 813
Strom, A. L., Steidel, C. C., Rudie, G. C., Trainor, R. F., & Pettini, M. 2018, ApJ, 868, 117
Tang, M., Stark, D. P., Chevallard, J., & Charlot, S. 2019, MNRAS, 489, 2572
Vanzella, E., Giavalisco, M., Inoue, A. K., et al. 2010, ApJ, 725, 1011
Vasei, K., Siana, B., Shapley, A. E., et al. 2016, ApJ, 831, 38
Weisz, D. R., Johnson, B. D., Johnson, L. C., et al. 2012, ApJ, 744, 44
Wise, J. H., & Cen, R. 2009, ApJ, 693, 984
Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, MNRAS, 442, 2560